



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/080,945	02/22/2002	Jane Wen Chang	11646-011001	6493

26161 7590 07/08/2004

FISH & RICHARDSON PC  
225 FRANKLIN ST  
BOSTON, MA 02110

EXAMINER
----------

WASSUM, LUKE S

ART UNIT	PAPER NUMBER
----------	--------------

2177

DATE MAILED: 07/08/2004

3

Please find below and/or attached an Office communication concerning this application or proceeding.

*[Handwritten signature]*

**Office Action Summary**

Application No.

10/080,945

Applicant(s)

CHANG ET AL.

Examiner

Luke S. Wassum

Art Unit

2177

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

- 1) ☒ Responsive to communication(s) filed on 22 February 2002.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

- 4) ☒ Claim(s) 1-20 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-20 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

**Application Papers**

- 9) ☒ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 22 February 2002 is/are: a) ☐ accepted or b) ☒ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some \* c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
  2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

- |  |   |
|--|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892)  | 4) <input type="checkbox"/> Interview Summary (PTO-413)<br>Paper No(s)/Mail Date. _____ |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)   | 5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152)             |
| 3) <input checked="" type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)<br>Paper No(s)/Mail Date <u>2</u> . | 6) <input type="checkbox"/> Other: _____  |

## DETAILED ACTION

### *The Invention*

1. The claimed invention is a method of document retrieval including assigning concept labels to documents contained in a collection according to grammar rules, receiving a query, converting the query to a query concept using the grammar rules, and mapping the query concept to a concept label.

### *Information Disclosure Statement*

2. The Applicants' Information Disclosure Statement, filed 25 February 2004, has been received and entered into the record. Since the Information Disclosure Statement complies with the provisions of MPEP § 609, the references cited therein have been considered by the examiner. See attached form PTO-1449.

### *Drawings*

3. The drawings are objected to as failing to comply with 37 CFR 1.84(p)(5) because they do not include the following reference character(s) mentioned in the description:

Numbers 48, 50 and 52, cited in the paragraph beginning on page 5, line 4 of the specification.

Corrected drawing sheets are required in reply to the Office action to avoid abandonment of the application. Any amended replacement drawing sheet should include all of the figures appearing on the immediate prior version of the sheet, even if only one figure is being amended. The replacement sheet(s) should be labeled "Replacement Sheet" in the page header (as per 37 CFR 1.84(c)) so as not to obstruct any portion of the drawing figures. If the changes are not accepted by

the examiner, the applicant will be notified and informed of any required corrective action in the next Office action. The objection to the drawings will not be held in abeyance.

4. The drawings are objected to because they fail to show necessary textual labels of features or symbols in Figs. 1-3 as described in the specification. For example, placing a label, "database", with element 28 of Fig. 1, would give the viewer necessary detail to fully understand this element at a glance. A descriptive textual label for each numbered element in these figures would be needed to better understand these figures without substantial analysis of the detailed specification. Any structural detail that is of sufficient importance to be described should be labeled in the drawing. Optionally, the applicant may wish to include a table next to the present figure to fulfill this requirement. See 37 CFR 1.84(n)(o), recited below:

"(n) Symbols. Graphical drawing symbols may be used for conventional elements when appropriate. The elements for which such symbols and labeled representations are used must be adequately identified in the specification. Known devices should be illustrated by symbols which have a universally recognized conventional meaning and are generally accepted in the art. Other symbols which are not universally recognized may be used, subject to approval by the Office, if they are not likely to be confused with existing conventional symbols, and if they are readily identifiable.

(o) Legends. Suitable descriptive legends may be used, or may be required by the Examiner, where necessary for understanding of the drawing, subject to approval by the Office. They should contain as few words as possible."

5. New corrected drawings are required in this application because the drawings, and particularly the handwritten labels, are of such poor quality as to render the drawings too difficult to read. Applicant is advised to employ the services of a competent patent draftsman outside the Office, as the U.S. Patent and Trademark Office no longer prepares new drawings. The corrected

Art Unit: 2177

drawings are required in reply to the Office action to avoid abandonment of the application. The requirement for corrected drawings will not be held in abeyance.

### *Specification*

6. The disclosure is objected to because of the following informalities:

On page 1, line 12, there is a typographical error: "...on the Web that is..." should be "...on the Web that *are*...";

On page 9, line 10, there is a typographical error: "The inflation is preformed..." should be "The inflation should be *performed*...".

Appropriate correction is required.

### *Claim Rejections - 35 USC § 102*

7. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

8. Claims 1-4, 8-10 and 13-20 are rejected under 35 U.S.C. 102(b) as being anticipated by Wical (U.S. Patent 6,038,560).

9. Regarding claim 1, Wical teaches a method of retrieving information comprising:

- a) assigning concept labels to documents contained in a collection (see disclosure that the content processing system processes a plurality of documents to identify themes and classifies the documents in categories in the knowledge base; see also disclosure of theme vectors, col. 6, line 34 through col. 7, line 2; see also col. 8, lines 1-11);
  - b) receiving a query (see col. 8, lines 14-15);
  - c) converting the query to a query concept (see disclosure of the identification of terminology having lexical, semantic or usage association with the query terms, col. 9, lines 34-53; see also identification of corresponding themes, col. 9, line 64 through col. 10, line 4); and
  - d) mapping the query concept to a concept label (see disclosure of identification of a category associated with query theme, col. 10, lines 1-4; see also col. 17, lines 27-61).
10. Regarding claim 13, Wical teaches a method of document retrieval as claimed, comprising:
- a) assigning concept labels to documents contained in a collection according to grammar rules (see disclosure that the content processing system processes a plurality of documents to identify themes and classifies the documents in categories in the knowledge base; see also disclosure of theme vectors, col. 6, line 34 through col. 7, line 2; see also col. 8, lines 1-11; see col. 27, line 14 through col. 28, line 24 for disclosure of grammar rules);
  - b) receiving a query (see col. 8, lines 14-15);
  - c) converting the query to a query concept using the grammar rules (see disclosure of the identification of terminology having lexical, semantic or usage association with the query terms, col. 9, lines 34-53; see also identification of corresponding themes, col. 9,

line 64 through col. 10, line 4; see col. 27, line 14 through col. 28, line 24 for disclosure of grammar rules); and

- d) mapping the query concept to a concept label (see disclosure of identification of a category associated with query theme, col. 10, lines 1-4; see also col. 17, lines 27-61).

11. Regarding claim 17, Wical teaches a computer program residing on a computer-readable medium as claimed, comprising instructions for causing a processor to:

- a) assign concept labels to documents contained in a collection (see disclosure that the content processing system processes a plurality of documents to identify themes and classifies the documents in categories in the knowledge base; see also disclosure of theme vectors, col. 6, line 34 through col. 7, line 2; see also col. 8, lines 1-11);
- b) receive a query (see col. 8, lines 14-15);
- c) convert the query to a query concept (see disclosure of the identification of terminology having lexical, semantic or usage association with the query terms, col. 9, lines 34-53; see also identification of corresponding themes, col. 9, line 64 through col. 10, line 4); and
- d) map the query concept to a concept label (see disclosure of identification of a category associated with query theme, col. 10, lines 1-4; see also col. 17, lines 27-61).

12. Regarding claim 19, Wical teaches a computer program residing on a computer-readable medium as claimed, comprising instructions for causing a processor to:

- a) assign concept labels to documents contained in a collection according to grammar rules (see disclosure that the content processing system processes a plurality of documents

to identify themes and classifies the documents in categories in the knowledge base;  
see also disclosure of theme vectors, col. 6, line 34 through col. 7, line 2; see also col.  
8, lines 1-11; see col. 27, line 14 through col. 28, line 24 for disclosure of grammar  
rules);

b) receive a query (see col. 8, lines 14-15);

c) convert the query to a query concept using the grammar rules (see disclosure of the  
identification of terminology having lexical, semantic or usage association with the  
query terms, col. 9, lines 34-53; see also identification of corresponding themes, col. 9,  
line 64 through col. 10, line 4; see col. 27, line 14 through col. 28, line 24 for  
disclosure of grammar rules); and

d) map the query concept to a concept label (see disclosure of identification of a category  
associated with query theme, col. 10, lines 1-4; see also col. 17, lines 27-61).

13. Regarding claims 2 and 14, **Wical** additionally teaches a method in which assigning  
comprises parsing the documents automatically with grammar rules (see col. 27, line 14 through col.  
28, line 24 for disclosure of grammar rules).

14. Regarding claim 3, **Wical** additionally teaches a method in which the concept label  
represents a general notion (see disclosure of identification of a category associated with query  
theme, col. 10, lines 1-4; see also col. 17, lines 27-61).

15. Regarding claims 4 and 15, **Wical** additionally teaches a method in which the query is a text  
query received from a user (see col. 1, lines 11-25).

16. Regarding claim 8, **Wical** additionally teaches a method in which converting comprises applying a store of grammar rules to the query (see col. 13, lines 7-23; see also col. 27, line 14 through col. 28, line 24 for a more specific disclosure of grammar rules).

17. Regarding claim 9, **Wical** additionally teaches a method in which the grammar rules map text to concepts (see col. 13, lines 7-23; see also col. 27, line 14 through col. 28, line 24 for a more specific disclosure of grammar rules).

18. Regarding claims 10, 16, 18 and 20, **Wical** additionally teaches a method and computer program further comprising instructions for causing a processor to:

- a) generate a list of the mapped query concepts (see Figures 10A and 10B; see also col. 9, lines 21-32; see also col. 13, lines 24-32); and
- b) display the list to a user on an input/output device (see Figures 10A and 10B; see also col. 9, lines 21-32; see also col. 13, lines 24-32).

### *Claim Rejections - 35 USC § 103*

19. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

- (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

Art Unit: 2177

20. The factual inquiries set forth in *Graham v. John Deere Co.*, 383 U.S. 1, 148 USPQ 459 (1966), that are applied for establishing a background for determining obviousness under 35 U.S.C. 103(a) are summarized as follows:

1. Determining the scope and contents of the prior art.
2. Ascertaining the differences between the prior art and the claims at issue.
3. Resolving the level of ordinary skill in the pertinent art.
4. Considering objective evidence present in the application indicating obviousness or nonobviousness.

21. This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

22. Claims 5-7, 11 and 12 are rejected under 35 U.S.C. 103(a) as being unpatentable over **Wical** (U.S. Patent 6,038,560) as applied to claims 1-4, 8-10 and 13-20 above, and further in view of **Braden-Harder et al.** (U.S. Patent 5,933,822).

23. Regarding claim 5, **Wical** teaches a method substantially as claimed.

**Wical** does not explicitly teach a method wherein the assignment of concept labels is performed by spidering the Internet and storing the location of discovered documents.

**Braden-Harder et al.**, however, teaches a method wherein the assignment of concept labels is performed by spidering the Internet and storing the location of discovered documents (see col. 1, line 49 through col. 2, line 30).

It would have been obvious to one of ordinary skill in the art at the time of the invention to spider the Internet to record a mapping of document features to concepts and document locations, since the Internet comprises a source of valuable information that is larger than any single conventional database (see col. 1, lines 36-48) and using this technique would significantly ease the task of retrieving information from the Internet (see col. 1, lines 49-52).

24. Regarding claim 6, **Braden-Harder et al.** additionally teaches a method wherein the documents are HyperText Markup Language (HTML) files (see disclosure of crawlers indexing the Internet, including HTML documents, col. 1, line 49 through col. 2, line 30).

25. Regarding claim 7, **Braden-Harder et al.** additionally teaches a method wherein the document location indicators are Universal Resource Identifiers (see disclosure that the document records typically include the URL associated with the document, col. 1, line 66 through col. 2, line 5).

26. Regarding claim 11, **Wical** teaches a method substantially as claimed.

Wical does not explicitly teach a method wherein the list of documents represents locations of documents, although the fact that documents could be retrieved via a network from a remote location is taught at col. 5, lines 35-41, which implies the storage of the claimed location information.

**Braden-Harder et al.**, however, teaches a method wherein the list of documents represents locations of documents (see disclosure that the document records typically include the URL associated with the document, col. 1, line 66 through col. 2, line 5).

It would have been obvious to one of ordinary skill in the art at the time of the invention to maintain the location of documents, since the Internet comprises a source of valuable information that is larger than any single conventional database (see col. 1, lines 36-48), and retrieval of information from the Internet requires the maintenance of location information.

27. Regarding claim 12, **Braden-Harder et al.** additionally teaches a method wherein the locations are Universal Resource Identifiers (see disclosure that the document records typically include the URL associated with the document, col. 1, line 66 through col. 2, line 5).

### *Conclusion*

28. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

**Turtle** (U.S. Patent 5,418,948) teaches a computer implemented process for creating a search query for an information retrieval system in which a database is provided containing a plurality of stopwords and phrases.

**Rubinstein** (U.S. Patent 5,721,897) teaches a method for identifying one of a plurality of documents stored in a computer-readable medium.

**Kirk et al.** (U.S. Patent 5,768,578) teaches an information retrieval system for retrieving information from a plurality of sources and for storing information source descriptions in a knowledge base.

**Wical** (U.S. Patent 5,940,821) teaches a knowledge base search and retrieval system which includes factual knowledge base queries and concept knowledge base queries.

**Tsourikov et al.** (U.S. Patent 6,167,370) teaches a method for semantically processing a user entered natural language request to identify and store linguistic subject-action-object (SAO) structures, and using such structures as key words/phrases to search local and web-based databases.

**Horowitz et al.** (U.S. Patent 6,236,987) teaches an information system that provides organizational and navigational aids to a user to facilitate exploration and analysis of a document collection.

**Steinkraus** (U.S. Patent 6,363,373) teaches a method of concept searching using a Boolean or keyword search engine.

**Mills** (U.S. Patent 6,466,940) teaches a system for automatically creating databases containing industry, service, product and subject classification data, contact data, geographic location data (CCG-data) and links to web pages from HTML, XML or SGML encoded web pages.

**Aggarwal et al.** (U.S. Patent 6,542,889) teaches a method of performing a conceptual similarity search.

**Kirkwood et al.** (U.S. Patent 6,665,662) teaches a technique for translating queries for related concepts in a database of concepts and relationships among concepts.

**Lin et al.** (U.S. Patent 6,675,159) teaches a concept-based indexing and search system which indexes collections of documents with ontology-based predicate structures through automated or human-assisted methods.

**Moore et al.** ("Web Page Categorization and Feature Selection Using Association Rule and Principal Component Clustering") teaches two new clustering algorithms that can effectively cluster documents, even in the presence of a very high dimensional feature space.

**Mase** ("Experiments on Automatic Web Page Categorization for IR Systems") teaches keyword-based Web page categorization.

**Berners-Lee et al.** ("The Semantic Web") teaches features of The Semantic Web.

The following reference, while not qualifying as prior art, is also of interest:

**Brasethvik et al.** ("A Conceptual Modeling Approach to Semantic Document Retrieval") teaches a prototype model-based classification and search tool.


Any inquiry concerning this communication or earlier communications from the examiner should be directed to Luke S. Wassum whose telephone number is 703-305-5706. The examiner can normally be reached on Monday-Friday 8:30-5:30, alternate Fridays off.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, John E. Breene can be reached on 703-305-9790. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

In addition, INFORMAL or DRAFT communications may be faxed directly to the examiner at 703-746-5658.

Customer Service for Tech Center 2100 can be reached during regular business hours at (703) 306-5631, or fax (703) 746-7240.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

  
Luke S. Wassum  
Art Unit 2177

lsw  
30 June 2004